TOWARDS QUANTUM MECHANICS FROM A THEORY OF EXPERIMENTS.

Inge S. Helland
Department of Mathematics,
University of Oslo,
P.O. Box 1053 Blindern,
N-0316 Oslo, Norway
E-mail: ingeh@math.uio.no

Abstract

Essential elements of quantum theory are derived from an epistemic point of view, i.e., the viewpoint that the theory has to do with what can be said about nature. This gives a relationship to statistical reasoning and to other areas of modelling and decision making. In particular, a quantum state can be defined from an epistemic point of view to consist of two elements: A (maximal) question about the value of some statistical parameter together with the answer to this question. Quantization itself can be approached from the point of view of model reduction under symmetry.

1 Introduction and background.

Recently, several theoretical physicists, most noteworthy Fuchs, [1, 2, 3] have argued for an epistemic interpretation of quantum mechanics. This means that quantum mechanics is taken to be concerned with what we can say about nature, not with how nature is. An implication argued for by Fuchs and others is that quantum states should be taken as subjective states of knowledge. I will formulate this in a slightly modified way in Section 11 below, but my basic point of view is that I fully agree with the epistemic interpretation of quantum theory, in fact I will argue that much of the theory can be derived in a natural way from this point of view.

To make this epistemic interpretation concrete, Fuchs and other authors like Schack and Caves [4, 5, 6, 7] have argued for and developed formally a Bayesian implementation of quantum mechanics.

An important remark now is that Bayesianism is a statistical concept. For a statistical formulation it is essential to distinguish between theoretical variables (statistical parameters) and observed variables (observations). This is already

seen from the simplest form of Bayes' formula, expressing conditional probabilities as

$$\mathrm{P}(H|D) = \frac{\mathrm{P}(H)\mathrm{P}(D|H)}{\mathrm{P}(D)},$$

where the data D are the observations, and H is some hypothesis about nature. These two items must be taken to have different ontological status. The hypothesis may often be expressed on the form, say, $a \leq \theta \leq b$ or $\lambda = k$, where θ and λ are statistical parameters. Thus the parameters are model concepts, words though which one can express hypotheses.

The distinction between observation and parameters is routinely implemented in biology, medicine, social sciences and other fields where the statistical modeling concept is being applied, but unfortunatly it has no tradition in theoretical physics. On the other hand, every experimental physicist knows that a measurement apparatus has an uncertainty, and that the result of the measurement process thus gives an uncertain estimate of a theoretical value. Following this line of thought, the distinction between theoretical statistical parameter and observed quantity becomes important also in physics. I will argue that it also is important in a new way to arrive at the foundation of quantum mechanics. This new way represents a radical departure from the ordinary way of deriving the quantum formalism. It is also a departure from ordinary statistical theory, since the parameter concept from statistics is extended somewhat.

In theoretical statistics, parameters are usually associated with an infinite population. In this paper, a parameter will denote anything which is unknown in the sense that it is relevant to request its value through experiment, see the examples below. Both in physics and in statistics one may have to consider technical errors, so that observations are modelled by probability distributions, given the parameters. In classical statistics this, possibly together with Bayesian priors, is the only basis for inference. Later in this paper we will show that we in addition, related to quantum mechanics, can make symmetry assumptions so that inference can be made from one experiment to another. In fact, this is our basis for deriving essential parts of quantum mechanics.

Through the years, several authors have tried to derive the formalism of quantum theory from more intuitive axioms. A list is given in the paper [8] by Hardy. Much of this list is from the quantum logic tradition, but it also includes the development of Mackey [9] using symmetry groups and the axiomatization of Accardi [10] related to Schwinger algebras. Further discussion about background papers for quantum mechanics is given in Helland [11].

The recent series of papers by Hardy [8, 12, 13] must also be added to the background list. In fact they give essential parts of quantum theory from simple assumptions indeed. Though we use a completely different approach, our results can also be related to those of Hardy. One main difference is that we - in the ordinary statistical tradition - assume more general measurements than measurements of probabilities.

It is interesting that it has been shown by Schack [14] that Hardy's frequentist probabilities can be replaced by Bayesian probabilities . Our own interpretation of this is that both kinds of probabilities can be used in the foundation

of quantum mechanics. In our approach it is natural to use priors derived from symmetry groups. It can be shown (Helland [15]) that for such priors there is a close connection between the use of the frequentist and the Bayesian formulation of statistics, although they from a logical point of view are different. The present author has a rather non-dogmatic relation to these two statistical schools: both have their strength and their weaknesses, and depending upon the application, both may be useful. A more important issue is the distinction between statistical parameter and observation.

2 Observations, parameters and total parameters.

Consider a single medical patient at some fixed time t. Let him be given some definite treatment, and let λ be his expected lifetime at this time. This is what I will call a statistical parameter. It is unknown, but can be assessed by medical expertise. It has a value in the sense that there ultimately will be an observation x which can be taken as a definite estimate of λ.

Now extend the situation. Assume that there at this time is a choice between two different treatments 1 and 2. Let λ^1 be the expected lifetime of the patient under treatment 1, and let λ^2 be the expected lifetime under treatment 2. Then each of λ^1 and λ^2 are parameters and have a value: There are experiments under which the parameters can be estimated.

Consider now the vector $\phi = (\lambda^1, \lambda^2)$. This is not a parameter in the sense that I have defined it. It can never take a value, since there is no experiment under which it can be estimated. It is just a mathematical variable. For our purpose it is important that such mathematical variables can not be considered as hidden variables in a physical theory either, since they can not be assigned any value.

As a physical analogy, let $\psi = (\xi, \pi)$ be the theoretical values of the position and momentum at some given particle at a fixed time. There is no experiment under which ψ can be given a value, but nevertheless one can consider ψ as a mathematical quantity.

Matematical quantities like ϕ and ψ will from now on be called *total parameters*. They can not be assigned any value, but nevertheless, operations like group transformations are applicable to the total parameters. Time scale transformations are relevant for ϕ, and the Galilei group together with translations proves the point for ψ.

Even though total parameters are incapable of taking any physical value, they can nevertheless be useful in contemplating models of physical and other phenomena.

3 The triangle-in-a-sphere example.

Let us consider an isolated part of reality consisting of a non-transparent sphere with three equidistant windows at the equator: 1, 2 and 3. We don't know what is inside the sphere, but we have a model saying that there is a rotating equilateral triangle there with corners A, B and C on the sphere. In any case, we have something inside which in a similar symmetrical way may produce the letters A, B or C close to the windows.

We assume three possible experiments here, but also assume that there is a mechanism such that these experiments are mutually exclusive: Look through window 1, through window 2 or through window 3. Let λ^a be the triangle corner closest to window a if this experiment is chosen ($a = 1, 2, 3$). There may be some uncertainty in the experiment which may be modeled by an ordinary statistical model, but it is assumed that λ^a is the only parameter that can be estimated in experiment a.

We now adapt the rotating triangle as our model, and let ϕ be the hypothetical position of this triangle inside the sphere. From our point of view this is a total parameter. It can not take any value that we can assess under the constraints described in the previous section. It may not exist physically at all; the reality inside the sphere may be richer than just a triangle; we have no possibility to know. Nevertheless, a transformation group G can in a natural way be associated with ϕ as a mathematical quantity, namely the rotation group.

This simple system satisfies the following assertions:

- For each a we have $\lambda^a = \lambda^a(\phi)$ - a valid parameter (cp. also the examples of Section 2, where a similar relation holds).

- For fixed a there is a subgroup G^a on which the following holds:

$$\lambda^a(\phi_1) = \lambda^a(\phi_2) \ implies \ \lambda^a(\phi_1 g) = \lambda^a(\phi_2 g) \ for \ g \in G^a.$$

(It is a general theorem that such a group exists, but it may be trivial. In the present case G^a can be taken as the group of rotations along the equator. The main point is: The subgroup G^a induces a transformation on the parameter λ^a.) (Throughout this paper, group operations are written to the right.)

- For any pair a, b there is a $g_{ab} \in G$ which transforms λ^a to λ^b. We have $g_{ac} = g_{ab} g_{bc}$.

- Each parameter λ^a takes only a finite number of values.

4 General assumptions.

Let us for some physical or other situation adopt the following assumptions. These are not artificial assumptions; they are motivated by examples such as the triangle-in-a-sphere.

- There is a set $\mathcal{A}$ of mutually exclusive experiments that can be done on some unit(s). The parameter of experiment a is λ^a.

- Define a total parameter ϕ such that each $\lambda^a = \lambda^a(\phi)$. Assume that there is a group G acting upon ϕ. Let $\Phi = \{\phi\}$. Note that we do not make any

physical or other invariance assumption at this point, only that it is possible and natural to define a transformation group G on Φ.

- There is a nontrivial subgroup G^a which leads to a transformation of λ^a, that is, one can define $\lambda^a g(\phi) = \lambda^a(\phi g)$ uniquely when $g \in G^a$.

- For any pair a, b there is a $g_{ab} \in G$ which transforms λ^a to λ^b, that is $\lambda^b(\phi) = \lambda^a(\phi g_{ab})$.

- We have $g_{ac} = g_{ab} g_{bc}$.

- Each parameter λ^a takes only a finite number of values.

- $\Phi = \{\phi\}$ is locally compact.

- The transformation group G as an action upon the total parameter space Φ has a right invariant measure ν on this space.

- The group G is compact (probably much can be extended to the non-compact case).

- The subgroups G^a generate G.

The last assumption is the only one that does not hold in the triangle-in-a-sphere example. This means that the theory below has to be modified somewhat for this example. (The purpose of this specific assumption is to find a general, natural representation for G; see Theorem 1 below. For the triangle case, the group G can be taken as the group of permutation of 3 elements, which has a well known 2-dimensional representation. The transition probabilities in this example are trivial, however.)

We turn to an example where all the assumptions above are satisfied.

5 A qubit model.

Consider an electron, and model its spin by a vector ϕ, a total parameter, i.e., just a mathematical variable connected to an abstract model. In the usual way of modelling, let the direction of ϕ be the spinning axis, and let the norm $\|\phi\|$ give the spinning speed. Let G be the group of rotations applied to ϕ.

To choose an experiment, choose some direction a in space. Define the parameter of this experiment, taking the values ± 1 as $\lambda(\phi) = \mathrm{sign}(a \cdot \phi)$. The real experiment will be of Stern-Gerlach type and may measure λ with some error. This error may be modeled by some ordinary statistical model, that is, a probability model for the data, given the value of λ, but we do not have to discuss this aspect.

It is easy to go through all the assumptions of the previous section and verify that they hold for this example. The subgroup G^a is the group of rotations around a plus a 180^{o} reflection around an axis perpendicular to a.

It may also be useful to consider the construction of λ as a two step process, where we first look upon the component of ϕ along the direction a, i.e., $\theta = \|\phi\| \cos(a, \phi)$ as a possible parameter. This also leads to the same subgroup G^a. But this group is highly non-transitive upon the range of θ, and λ may be taken as a model reduction (see below) corresponding to exactly one orbit of this group.

6 A Hilbert space for experiment a.

Make the general assumptions above from now on. Take as a basic space $L = L^2(\Phi, \nu)$, considering complex functions. Let $\mathbf{H}^a$ be the subspace of L consisting of functions of $\lambda^a(\phi)$:

$$\mathbf{H}^a = \{f \in L : f(\phi) = \tilde{f}(\lambda^a(\phi))\}.$$

Since λ^a only takes a finite number of values, this will be a finite-dimensional space, and hence a Hilbert space. The theory can be generalised to infinite-dimensional spaces, but this will require more technicalities.

$\mathbf{H}^a$ is an invariant space under the right regular representation of the subgroup G^a.

Let the possible values for λ^a be λ_k $(k = 1, ..., n)$. Then a basis for $\mathbf{H}^a$ is given by the indicators

$$f_k^a(\phi) = I_{\lambda_k}(\lambda^a(\phi)); (k = 1, ..., n).$$

These are eigenvectors of the trivial operator S^a defined on $\mathbf{H}^a$ by

$$S^a \tilde{f}(\lambda^a(\phi)) = \lambda^a(\phi)\tilde{f}(\lambda^a(\phi)).$$

7 A common Hilbert space.

The different spaces $\mathbf{H}^a$ are related through

$$\mathbf{H}^b = U(g_{ab})\mathbf{H}^a,$$

where $U(\cdot)$ is the right regular group representation on $L^2(\Phi, \nu)$. (Helland [11].)

To begin with, fix one c and take $\mathbf{H} = \mathbf{H}^c$. Then, defining $E^a = U(g_{ca})$, we have the connections $\mathbf{H}^a = E^a\mathbf{H}$.

Recall that we have assumed that the full group G is generated by the subgroups G^a.

Theorem 1.
$\mathbf{H}$ *is an invariant space for* G *under the representation exemplified by*

$$W(g_1 g_2 g_3) = E^{a\dagger}U(g_1)E^a E^{b\dagger}U(g_2)E^b E^{c\dagger}U(g_3)E^c,$$

if $g_1 \in G^a$, $g_2 \in G^b$ and $g_3 \in G^c$.

(See Helland [11].) The point is that $\mathbf{H}$ always can be regarded as an invariant space of *some* representation of the full group G. If necessary, one can consider a subrepresentation of $W(\cdot)$. In a quantum mechanical setting, irreducible subrepresentations are related to superselection rules. In general one can have $\mathbf{H} = \mathbf{H}_1 \oplus \mathbf{H}_2 \oplus \ldots$, where the components $\mathbf{H}_i$ are invariant spaces under an irreducible representation. For simplicity we will assume one component in the following.

It is well known that every representation of a compact group by a unitary transformation can be made into a subrepresentation of the regular representation. Now make such a unitary change of $\mathbf{H}$ and $W(\cdot)$ such that $W(\cdot)$ is a subrepresentation of the regular representation on L. During this, each f_j^c changes to some unit vector which we will call $|c, j>$ to conform to the ordinary quantum mechanical notation, and S^c changes to an operator T^c. Note that all Hilbert spaces of the same dimension are unitarily equivalent, so that it is of no importance that we use a particular construction here.

Define now the fundamental state vectors $|a, k>$ by $W(g_{ca})|c, k>$ and the corresponding operators T^a by $T^a = W(g_{ca})T^c W(g_{ac})$.

Proposition 1.
The vectors $|a, k>$ are eigenvectors of T^a with eigenvalues $\lambda_k = \lambda_k = \lambda_k^c$.

One can show in general [11] that in the present setting the set of eigenvalues is the same for the operators of all experiments. Thus one does not obtain the most general quantummechanical Hilbert space here, but the framework includes qubits, higher spins, sets of particle with spins and the most common entanglement case.

Note that the simple indicator functions $f_k^a(\cdot) = I_{\lambda_k}(\lambda^a(\cdot))$ are related by $f_k^a = U(g_{ca})f_k^c$, and that $|c, k>$ is just a fixed unitary transformation of f_k^c. This, together with the similar relation $|a, k >= W(g_{ac})|c, k >= U(g_{ac})|c, k >$, may be taken to connect $|a, k>$ to the statement $\lambda^a = \lambda_k$, or, more precisely, motivate the following definition:

Definition 1.
The state vector $|a, k>$ is per definition taken to represent two elements:
1. A question: What is the value of λ^a? This corresponds to a choice of experiment a, and, more specifically, to a perfect experiment.
2. An answer: $\lambda^a = \lambda_k$.

Note that the vectors $|a, k>$ as defined above always give an orthonormal basis for $\mathbf{H}$. The way we have constructed the Hilbert space here, it has always the same dimension as the number of distinct values of λ^a. This means that the set of answers is maximal, and, relative to the operator T^a, the eigenvalues λ_k are non-degenerate. In other cases, it might be useful to consider λ^a as a subparameter of another parameter θ^a, and to let this latter parameter define the Hilbert space. Then the parameter values λ_k will be degenerate eigenvalues of the corresponding operator. For the non-degenerate case, in the above formulation, a parameter λ^a is put in one-to-one correspondence with a resolution of the identity in $\mathbf{H}$.

In conventional quantum mechanics, every state vector is the eigenvector of *some* operator, in fact many such. What we assume here is essentially that we always can find such an operator which is of physical relevance. In fact, via the parameter we go directly to the quantity of physical relevance, and

give a nonformal, epistemic definition of a state. In most cases, the question-and-answer is directly related to experiments, but we can also have indirect assessments of states, as in the EPR experiment (see below).

Since there is an arbitrary phase factor in the fixed unitary transformation use above, there is an arbitrary phase factor in $|a, k>$, so it is even more precise to say that the two elements above are represented by the one-dimensional projector $|a, k >< a, k|$. The results of most real experiment also involve experimental uncertainty, and through this, a statistical model. The final result is then given by the a posteriori Bayesian probabilities π_k connected to each parameter value of the chosen experiment and thus to a density matrix

$$\rho = \sum_k \pi_k |a, k >< a, k|.$$

Such a density matrix always describes the state whenever there is an uncertainty π_k related to the answer to the question a. In particular, the probabilities π_k can also be prior probabilities.

As constructed above, there is also an operator T^a on $\mathbf{H}$ connected to the parameter itself, namely,

$$T^a = \sum_k \lambda_k |a, k >< a, k|.$$

Note that in our approach, the parameter, i.e., the question, comes first; the operator is a derived quantity.

The remaining problem is to characterise those vectors in $\mathbf{H}$ that can be taken as state vectors in the above sense. In ordinary quantum mechanics (without superselection) we would want all unit vectors to be state vectors, but this will require further assumptions on the set of experiments, briefly, it should be rich enough. We have the following result (Helland [16]):

Theorem 2.

a) Every element of the group G can be written as $g = g^c g_{cb}$ for some $g^c \in G^c$.

b) Fix $|0 >= |c, j > \in \mathbf{H}$. Then every state vector $|a, k>$ can be written as $|a, k >= W(g)|0>$ for some $g \in G$, that is, as a generalised coherent state (GCS), and all such vectors are state vectors.

Open problem.

Can one find simple conditions under which the generalised coherent vectors (with phase changes) give all the unit vectors in $\mathbf{H}$?

In the qubit case one can use a Bloch sphere argument to show that indeed the state vectors constitute all the unit vectors in the Hilbert space.

8 Transition probabilities.

Assume that we start in some state, say, given by the fact that a perfect measurement of λ^a has resulted in the value λ_k. This corresponds to a state vector $|a,k>$.

We now ask ourselves: What if we do a new perfect measurement of another parameter λ^b, what is the the probability of getting $\lambda^b = \lambda_i$, say?

The answer is given by Born's formula:

$$P(\lambda^b = \lambda_i | \lambda^a = \lambda_k) = | < a,k | b,i > |^2.$$

As is well known, this formula can be taken as the starting point for large parts of the quantum formalism.

9 On the proof of Born's formula.

Busch [17] introduced the concept of an effect

$$E = \sum_i p_i |i >< i|,$$

where $\{|i>\}$ is an orthonormal basis of the Hilbert space, and p_i are probabilities between 0 and 1. These can be given a statistical interpretation: Let $\{\lambda_i\}$ be the possible parameter values of some chosen experiment b, and let $|i> = |b,i>$ be the corresponding pure state vectors, indicating that $\lambda^b = \lambda_i$. Let the model for the whole experiment, assuming discrete observations and discrete parameter be given by $p_i = p(x|\lambda_i)$. Then the operator E characterizes the whole experiment, including choice of parameter.

A transition probability from a fixed initial state to an arbitrary effect E is called a generalized probability if

$$P(E_1 + E_2 + \ldots) = P(E_1) + P(E_2) + \ldots$$

whenever $E_1 + E_2 + \ldots$ is an effect.

A main result of Busch [17] is the following variant of Gleason's theorem: *Any generalized probability on effects is of the form $P(E) = \mathrm{tr}(\rho E)$ for some density matrix ρ.*

Using this result, Born's formula can be relatively easily proved if we can show that the transition formula is a generalized probability. This in turn is proved in our setting by a statistical argument using the following assumptions, which then are our assumptions behind Born's formula and the quantummechanical results which follow in the next section:

(i) The transition probabilities $P(\lambda^b = \lambda_i | \lambda^a = \lambda_k)$ exist.
(ii) $P(\lambda^a = \lambda_k | \lambda^a = \lambda_k) = 1$.
(iii) For all a, b, c we have that $\mu(\phi) = \lambda^a(\phi g_{bc})$ is a valid parameter.

(iv) For all a, b, c, k, i we have that $P(\lambda^b(\phi) = \lambda_i | \lambda^a(\phi) = \lambda_k) = P(\lambda^b(\phi g_{bc}) = \lambda_i | \lambda^a(\phi g_{bc}) = \lambda_k)$.

The proof is carried out in Helland [11] by proving the relation

$$P(\frac{1}{2}(E_1 + E_2)) = \frac{1}{2}P(E_1) + \frac{1}{2}P(E_2)$$

when the arguments involved are effects. The main steps in the proof is first to rotate the experiment E_2 in such a way that all final state vectors agree with those of experiment E_1, then select between these experiments with probability $\frac{1}{2}$ each, and finally to rotate back.

10 About quantum mechanics from Born's formula.

From Born's formula one easily deduces standard formulas like

$$E(\lambda^b | \lambda^a = \lambda_k) = < a, k | T^b | a, k >,$$

where $T^b = \sum_j \lambda_j |b, j >< b, j|$, and so on.

This gives the essential elements of ordinary quantum mechanics. The elements that are lacking, Planck's constant and the Schrödinger equation, are discussed in Helland [18].

Connected to an experiment, Bayesian updating is done in the natural way; there is no mysterious collapse of a wave package.

To give some details, in ordinary statistical theory, the result of an experiment with parameter λ^b is modelled through a probability measure $P(dy | \lambda^b = \lambda_j)$ indexed by the value of the parameter. From this one can define an operator valued measure by

$$M(dy) = \sum_j P(dy | \lambda^b = \lambda_j) |b, j >< b, j|.$$

Then, from Born's formula, given some initial state $\lambda^a = \lambda_k$, stating that the question about the value of the maximal parameter λ^a is answered by λ_k, the probability distribution of the result of experiment b is given by

$$P[dy | \lambda^a = \lambda_k] = < a, k | M(dy) | a, k > .$$

More generally, if the initial state is given by a density matrix ρ, expressing uncertainty about the value of λ^a, we get $P[dy] = \mathrm{tr}[\rho M(dy)]$.

Disregard for simplicity now measurement errors, and assume again that the starting state is given by the density matrix

$$\rho = \sum_k \pi_k |a, k >< a, k|,$$

so that the probability is π_k that the initial state is given by $\lambda^a = \lambda_k$. Let us then measure another parameter λ^b with the corresponding unit vectors $|b, j>$. Then from Born's formula, the probability of getting the result λ_j is

$$\kappa_j^b = \sum_k \pi_k |<b, j|a, k>|^2,$$

so that the state after measurement but before the measurement value is found is given by

$$\sum_j \kappa_j^b |b, j><b, j| = \sum_j <b, j|\rho|b, j> |b, j><b, j|.$$

After the measurement value λ_j is found, the state is simply $|b, j><b, j|$.

We will not discuss here further general developments of quantum theory from the parameter starting point, but it should be reasonably clear in which direction the further arguments go.

11 Bohm's version of the EPR experiment. Entanglement.

Assume that the spin of two electrons which collide are given by the total parameters ϕ_1 and ϕ_2. Furthermore, assume that the state after the collision is such that $\phi_1 = -\phi_2$, so that the joint system can be modelled by one of these total parameters.

At some distance, the spin component in direction a for electron 1 is given by the parameter $\lambda^a(\phi_1)$, while the spin component in direction b for electron 2 is given by the parameter $\mu^b(\phi_2) = -\mu^b(\phi_1)$.

It results from theory (Born's formula) (and from experiments) that

$$E(\lambda^a \mu^b) = -a \cdot b.$$

In particular it follows from this that Bell's inequalities are violated.

The interpretation of this in our setting is as follows: Bell's inequalities require local realism. Here we have lack of *realism*: The total parameters ϕ_1 and ϕ_2 are just mathematical quantities; they do not take any value. Nevertheless, the system is bound together by a common value of the total parameter, a fact which may be related to the common history of the two parts. This relationship is given by model quantities, however; the two parts have nothing physically in common. But this common model quantity is enough to imply that the assumptions behind Bell's inequalities break down.

In Helland [18] there is a discussion of Bell's inequality using the statistical conditionality principle: Every inference should be conditioned upon the experiment which actually is performed. If we stick to this principle, the whole situation seems to be less paradoxical. In Helland [20] there is a tentative discussion to the effect that related situations are not impossible to imagine in

the macroscopic world, either. The concept of a total parameter is very central here.

The above experiment is much discussed in the quantum mechanical literature. Among other things it is used by Caves et al [5] to argue that, since the result $\lambda^a = \lambda_k$ for experiment a at site 1 implies that the state at the other site must be given by $\mu^a = -\lambda_k$, then these states must be just subjective states of knowledge. In the spirit of the present paper, a state is given by a question plus an answer, and it is the question posed which is subjective, given by the observer at site 1. The answer is given by the situation. But the resulting state is a state of knowledge.

12 Model reduction and quantization.

Note that we have introduced parameters as the words that we can use to describe nature. In applied statistics it is in many cases imperative to use simple models, i.e., the parameter space should have as simple structure as possible. We speak about a wide model with parameter θ and a narrow model with another parameter λ, and then in most cases $\lambda = f(\theta)$ for some function f and/or the range of θ is restricted to some subspace (see [19] and references there). In [15] it is argued that this subspace should be an orbit or a set of orbits of the relevant group.

These ideas can also be extended to total parameters. If a wide parameter θ^a corresponds to experiment a, then the most wide total parameter imaginable is given by

$$\pi = \times_a \theta^a,$$

a cartesian product. A more useful parameter may be given by $\phi = f(\pi)$, where f is a function which is *natural*, i.e., such that $f(\pi_1) = f(\pi_2)$ implies $f(\pi_1 g) = f(\pi_2 g)$ for all $g \in G$, and where π is restricted to a subspace Ψ, an orbit or a set of orbits of G.

A first step towards quantization will then be to consider the parameter values μ^a for which

$$\{\pi \in \Psi : \theta^a(\pi) = \mu^a\}$$

is nonempty. As in the rest of this paper, this induces a subgroup G^a acting upon the parameter μ^a. Model reduction as above on these single parameters may then be applied to give a final reduced parameter λ^a. Again it is required that the range of λ^a should constitute an orbit or a set of orbits of G^a.

For an application of these ideas to the qubit model, see [11].

13 Conclusions.

- Any theory takes a set of assumptions as the point of departure; here the assumptions may be related to fairly simple examples. In this way it is a non-formal theory.

- The theory seems to point at a connection between quantum mechanics and statistics. But choice of experiment among several complementary ones, and also symmetry assumptions, are needed to derive quantum mechanics.

- A large subset of the quantum mechanical state vectors, under certain assumptions all the state vectors, may be interpreted as (maximal) questions plus answers. Note that a question can consist of several partial questions, corresponding to commuting operators in the ordinary formalism.

- Symmetry considerations are important both in the foundation of quantum mechanics and in the foundation of statistics.

- There are no hidden variables, but hidden total parameters which are connected to mathematical models, and which never take any physical value.

- The natural implied interpretation is epistemic rather than ontological. In fact, the whole approach is epistemic or information theoretical, and can to a large extent be said to meet the imperatives for a new programme set forth in Fuchs [3].

- There is of course more to be done in the development of this theory. A natural continuation is to look at continuous parameters, which seem to be possible to describe in the same setting, but which may require more technicalities.

References

[1] C. A. Fuchs, quant-ph/0205039 (2002).

[2] C. A. Fuchs, *Quantum States: W.H.A.T?* http://netlib.bell-labs.com/who/cafuchs .

[3] C. A. Fuchs, *J. Mod. Opt.***50**, 987-1023 (2003).

[4] R. Schack, T. A. Brun and C. M. Caves, *Phys. Rev. A* **64**, 014305 (2001).

[5] C. M. Caves, C. A. Fuchs and R. Schack, *Phys. Rev. A.* **65**, 022305 (2002).

[6] C. M. Caves, C. A. Fuchs and R. Schack, quant-ph/02104088 (2001).

[7] C. A. Fuchs and R. Schack, In: M. G. A. Paris and J. Rehacek [Ed.] *Quantum Estimation Theory* Springer (2004)

[8] L. Hardy, quant-ph/0101012 (2001).

[9] G. W. Mackey, *The Mathematical Foundations of Quantum Mechanics* Benjamin, New York.

[10] L. Accardi, *Il Nuovo Cimento* **110B**, 685 (1995).

[11] I. S. Helland, to appear in *Ann. Statistics* **34**, (2006).

[12] L. Hardy, quant-ph/0111068 (2001).

[13] L. Hardy, In: J. D. Barrow, P. C. Davies and C. L. Harper, Jr [Ed.] *Science and Ultimate Reality* Cambridge University Press, 2004.

[14] R. Schack, quant-ph/0210017 (2002).

[15] I. S. Helland, *Internat. Statist. Review* **72**, 409 (2004).

[16] I. S. Helland, quant-ph/0507200 (2005).

[17] P. Busch, *Phys. Rev. Letters* **91**, 12 (2003).

[18] I. S. Helland, unpublished.

[19] N. L. Hjort and G. Claeskens, *J. Amer. Statist. Assn.* **98**, 879 (2003).

[20] I. S. Helland, quant-ph/0507205 (2005)